RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _10/574,902_

Source: _1FWP_

Date Processed by STIC: _4/26/06_

ENTERED

IFWP

RAW SEQUENCE LISTING   DATE: 04/26/2006
PATENT APPLICATION: US/10/574,902   TIME: 07:34:04

Input Set : **F:\Final Sequence List-13987-00003-US.txt**
Output Set: **N:\CRF4\04262006\J574902.raw**

```
 3  <110> APPLICANT: Cirpus, Petra
 4        Oswald, Oliver
 5        Lerchl, Jens
 6        Martin, William Frank
 7        Hoffmeister, Meike
 9  <120> TITLE OF INVENTION: Trans-2-enoyl-CoA reductase gene of Euglena gracilis
11  <130> FILE REFERENCE: 13987-00003-US
13  <140> CURRENT APPLICATION NUMBER: US/10/574,902
13  <141> CURRENT FILING DATE: 2006-04-07
13  <150> PRIOR APPLICATION NUMBER: PCT/EP2004/011294
14  <151> PRIOR FILING DATE: 2004-10-08
16  <150> PRIOR APPLICATION NUMBER: EP 03022783.9
17  <151> PRIOR FILING DATE: 2003-10-10
19  <150> PRIOR APPLICATION NUMBER: EP 04007051.8
20  <151> PRIOR FILING DATE: 2004-03-24
22  <160> NUMBER OF SEQ ID NOS: 55
24  <170> SOFTWARE: PatentIn version 3.3
27  <210> SEQ ID NO: 1
28  <211> LENGTH: 1620
29  <212> TYPE: DNA
30  <213> ORGANISM: Euglena gracilis
32  <220> FEATURE:
33  <221> NAME/KEY: CDS
34  <222> LOCATION: (1)..(1620)
36  <400> SEQUENCE: 1
37  atg tcg tgc ccc gcc tcg ccg tct gct gcc gtg gtg tct gcc ggc gcc        48
38  Met Ser Cys Pro Ala Ser Pro Ser Ala Ala Val Val Ser Ala Gly Ala
39  1               5                   10                  15
41  ctc tgc ctg tgc gtg gca acg gta ttg ttg gcg act gga tcc aac ccc        96
42  Leu Cys Leu Cys Val Ala Thr Val Leu Leu Ala Thr Gly Ser Asn Pro
43              20                  25                  30
45  acc gcc ctg tcc act gct tcc act cgc tct ccg acc tca ctg gtc cgt       144
46  Thr Ala Leu Ser Thr Ala Ser Thr Arg Ser Pro Thr Ser Leu Val Arg
47          35                  40                  45
49  ggg gtg gac agg ggc ttg atg agg cca acc act gca gcg gct ctg acg       192
50  Gly Val Asp Arg Gly Leu Met Arg Pro Thr Thr Ala Ala Ala Leu Thr
51      50                  55                  60
53  aca atg aga gag gtg ccc cag atg gct gag gga ttt tca ggc gaa gcc       240
54  Thr Met Arg Glu Val Pro Gln Met Ala Glu Gly Phe Ser Gly Glu Ala
55  65                  70                  75                  80
57  acg tct gca tgg gcc gcc gcg ggg ccg cag tgg gcg gcg ccg ctc gtg       288
58  Thr Ser Ala Trp Ala Ala Ala Gly Pro Gln Trp Ala Ala Pro Leu Val
59              85                  90                  95
```

RAW SEQUENCE LISTING   DATE: 04/26/2006
PATENT APPLICATION: US/10/574,902   TIME: 07:34:04

Input Set : F:\Final Sequence List-13987-00003-US.txt
Output Set: N:\CRF4\04262006\J574902.raw

```
 61 gcc gcg gcc tcc tcc gca ctg gcg ctg tgg tgg tgg gcc gcc cgg cgc      336
 62 Ala Ala Ala Ser Ser Ala Leu Ala Leu Trp Trp Trp Ala Ala Arg Arg
 63             100             105             110
 65 agc gtg cgg cgg ccg ctg gca gcg ctg gcg gag ctg ccc acc gcg gtc      384
 66 Ser Val Arg Arg Pro Leu Ala Ala Leu Ala Glu Leu Pro Thr Ala Val
 67         115             120             125
 69 acc cac ctg gcc ccc ccg atg gcg atg ttc acc acc aca gcg aag gtc      432
 70 Thr His Leu Ala Pro Pro Met Ala Met Phe Thr Thr Thr Ala Lys Val
 71     130             135             140
 73 atc cag ccc aag att cgt ggc ttc atc tgc acg acc acc cac ccg atc      480
 74 Ile Gln Pro Lys Ile Arg Gly Phe Ile Cys Thr Thr Thr His Pro Ile
 75 145             150             155             160
 77 ggc tgt gag aag cgg gtc cag gag gag atc gcg tac gcc cgt gcc cac      528
 78 Gly Cys Glu Lys Arg Val Gln Glu Glu Ile Ala Tyr Ala Arg Ala His
 79             165             170             175
 81 ccg ccc acc agc cct ggc ccg aag agg gtg ctg gtc atc ggc tgc agt      576
 82 Pro Pro Thr Ser Pro Gly Pro Lys Arg Val Leu Val Ile Gly Cys Ser
 83             180             185             190
 85 acc ggc tac ggg ctc tcc acc cgc atc acc gct gcc ttc ggc tac cag      624
 86 Thr Gly Tyr Gly Leu Ser Thr Arg Ile Thr Ala Ala Phe Gly Tyr Gln
 87         195             200             205
 89 gcc gcc acg ctg ggc gtg ttc ctg gcg ggc ccc ccg acg aag ggc cgc      672
 90 Ala Ala Thr Leu Gly Val Phe Leu Ala Gly Pro Pro Thr Lys Gly Arg
 91     210             215             220
 93 ccc gcc gcg gcg ggc tgg tac aac acc gtg gcg ttc gag aag gcc gcc      720
 94 Pro Ala Ala Ala Gly Trp Tyr Asn Thr Val Ala Phe Glu Lys Ala Ala
 95 225             230             235             240
 97 ctg gag gcc ggg ctg tac gcc cgg agc ctt aat ggc gac gcc ttc gac      768
 98 Leu Glu Ala Gly Leu Tyr Ala Arg Ser Leu Asn Gly Asp Ala Phe Asp
 99             245             250             255
101 tcc aca acg aag gcg cgg acg gtc gag gcg atc aag cgg gac ctc ggc      816
102 Ser Thr Thr Lys Ala Arg Thr Val Glu Ala Ile Lys Arg Asp Leu Gly
103             260             265             270
105 acg gtg gac ctc gtg gtg tac agc atc gcc gcc ccg aag cgg acg gac      864
106 Thr Val Asp Leu Val Val Tyr Ser Ile Ala Ala Pro Lys Arg Thr Asp
107         275             280             285
109 cct gcc acc ggc gtc ctc cac aag gcc tgc ctg aag ccc atc ggc gcc      912
110 Pro Ala Thr Gly Val Leu His Lys Ala Cys Leu Lys Pro Ile Gly Ala
111     290             295             300
113 acg tac acc aac cgc act gtg aac acc gac aag gcg gag gtg acc gac      960
114 Thr Tyr Thr Asn Arg Thr Val Asn Thr Asp Lys Ala Glu Val Thr Asp
115 305             310             315             320
117 gtc agc att gag ccg gcc tcc ccc gaa gag atc gcg gac acg gtg aag     1008
118 Val Ser Ile Glu Pro Ala Ser Pro Glu Glu Ile Ala Asp Thr Val Lys
119             325             330             335
121 gtg atg ggc ggg gag gac tgg gag ctc tgg atc cag gcg ctg tcg gag     1056
122 Val Met Gly Gly Glu Asp Trp Glu Leu Trp Ile Gln Ala Leu Ser Glu
123             340             345             350
125 gcc ggc gtg ctg gcg gag ggg gcc aag acg gtg gcg tac tcc tac atc     1104
```

RAW SEQUENCE LISTING   DATE: 04/26/2006
PATENT APPLICATION: US/10/574,902   TIME: 07:34:04

Input Set : F:\Final Sequence List-13987-00003-US.txt
Output Set: N:\CRF4\04262006\J574902.raw


```
126 Ala Gly Val Leu Ala Glu Gly Ala Lys Thr Val Ala Tyr Ser Tyr Ile
127         355                 360                 365
129 ggc ccc gag atg acg tgg cct gtc tac tgg tcc ggc acc atc ggg gag     1152
130 Gly Pro Glu Met Thr Trp Pro Val Tyr Trp Ser Gly Thr Ile Gly Glu
131         370                 375                 380
133 gcc aag aag gac gtg gag aag gct gcc aag cgc atc acg cag cag tac     1200
134 Ala Lys Lys Asp Val Glu Lys Ala Ala Lys Arg Ile Thr Gln Gln Tyr
135 385                 390                 395                 400
137 ggc tgc ccg gcg tac ccg gtg gtg gcc aag gcc ttg gtc acc cag gcc     1248
138 Gly Cys Pro Ala Tyr Pro Val Val Ala Lys Ala Leu Val Thr Gln Ala
139         405                 410                 415
141 agc tcc gcc atc ccg gtg gtg ccg ctc tac atc tgc ctg ctg tac cgc     1296
142 Ser Ser Ala Ile Pro Val Val Pro Leu Tyr Ile Cys Leu Leu Tyr Arg
143         420                 425                 430
145 gtt atg aag gag aag ggc acc cac gag ggc tgc atc gag cag atg gtg     1344
146 Val Met Lys Glu Lys Gly Thr His Glu Gly Cys Ile Glu Gln Met Val
147         435                 440                 445
149 cgg ctg ctc acc acg aag ctg tac ccc gag aac ggg gcc ccc atc gtc     1392
150 Arg Leu Leu Thr Thr Lys Leu Tyr Pro Glu Asn Gly Ala Pro Ile Val
151         450                 455                 460
153 gat gag gcc gga cgt gtg cgg gtg gat gac tgg gag atg gcg gag gat     1440
154 Asp Glu Ala Gly Arg Val Arg Val Asp Asp Trp Glu Met Ala Glu Asp
155 465                 470                 475                 480
157 gtg cag cag gct gtt aag gac ctc tgg agc cag gtg agc act gcc aac     1488
158 Val Gln Gln Ala Val Lys Asp Leu Trp Ser Gln Val Ser Thr Ala Asn
159         485                 490                 495
161 ctc aag gac atc tcc gac ttc gct ggg tat caa act gag ttc ctg cgg     1536
162 Leu Lys Asp Ile Ser Asp Phe Ala Gly Tyr Gln Thr Glu Phe Leu Arg
163         500                 505                 510
165 ctg ttc ggg ttc ggc att gac ggc gtg gac tac gac cag ccc gtg gac     1584
166 Leu Phe Gly Phe Gly Ile Asp Gly Val Asp Tyr Asp Gln Pro Val Asp
167         515                 520                 525
169 gtg gag gcg gac ctc ccc agt gct gcc cag cag tag                     1620
170 Val Glu Ala Asp Leu Pro Ser Ala Ala Gln Gln
171         530                 535
174 <210> SEQ ID NO: 2
175 <211> LENGTH: 539
176 <212> TYPE: PRT
177 <213> ORGANISM: Euglena gracilis
180 <400> SEQUENCE: 2
182 Met Ser Cys Pro Ala Ser Pro Ser Ala Ala Val Val Ser Ala Gly Ala
183 1                   5                   10                  15
186 Leu Cys Leu Cys Val Ala Thr Val Leu Leu Ala Thr Gly Ser Asn Pro
187         20                  25                  30
190 Thr Ala Leu Ser Thr Ala Ser Thr Arg Ser Pro Thr Ser Leu Val Arg
191         35                  40                  45
194 Gly Val Asp Arg Gly Leu Met Arg Pro Thr Thr Ala Ala Ala Leu Thr
195 50                  55                  60
198 Thr Met Arg Glu Val Pro Gln Met Ala Glu Gly Phe Ser Gly Glu Ala
```

```
199 65                        70                    75                          80
202 Thr Ser Ala Trp Ala Ala Ala Gly Pro Gln Trp Ala Ala Pro Leu Val
203              85                    90                    95
206 Ala Ala Ala Ser Ser Ala Leu Ala Leu Trp Trp Trp Ala Ala Arg Arg
207              100                   105                   110
210 Ser Val Arg Arg Pro Leu Ala Ala Leu Ala Glu Leu Pro Thr Ala Val
211          115                   120                   125
214 Thr His Leu Ala Pro Pro Met Ala Met Phe Thr Thr Thr Ala Lys Val
215      130                   135                   140
218 Ile Gln Pro Lys Ile Arg Gly Phe Ile Cys Thr Thr Thr His Pro Ile
219 145                   150                   155                   160
222 Gly Cys Glu Lys Arg Val Gln Glu Glu Ile Ala Tyr Ala Arg Ala His
223              165                   170                   175
226 Pro Pro Thr Ser Pro Gly Pro Lys Arg Val Leu Val Ile Gly Cys Ser
227              180                   185                   190
230 Thr Gly Tyr Gly Leu Ser Thr Arg Ile Thr Ala Ala Phe Gly Tyr Gln
231          195                   200                   205
234 Ala Ala Thr Leu Gly Val Phe Leu Ala Gly Pro Pro Thr Lys Gly Arg
235      210                   215                   220
238 Pro Ala Ala Ala Gly Trp Tyr Asn Thr Val Ala Phe Glu Lys Ala Ala
239 225                   230                   235                   240
242 Leu Glu Ala Gly Leu Tyr Ala Arg Ser Leu Asn Gly Asp Ala Phe Asp
243              245                   250                   255
246 Ser Thr Thr Lys Ala Arg Thr Val Glu Ala Ile Lys Arg Asp Leu Gly
247          260                   265                   270
250 Thr Val Asp Leu Val Val Tyr Ser Ile Ala Ala Pro Lys Arg Thr Asp
251          275                   280                   285
254 Pro Ala Thr Gly Val Leu His Lys Ala Cys Leu Lys Pro Ile Gly Ala
255      290                   295                   300
258 Thr Tyr Thr Asn Arg Thr Val Asn Thr Asp Lys Ala Glu Val Thr Asp
259 305                   310                   315                   320
262 Val Ser Ile Glu Pro Ala Ser Pro Glu Glu Ile Ala Asp Thr Val Lys
263              325                   330                   335
266 Val Met Gly Gly Glu Asp Trp Glu Leu Trp Ile Gln Ala Leu Ser Glu
267          340                 345                   350
270 Ala Gly Val Leu Ala Glu Gly Ala Lys Thr Val Ala Tyr Ser Tyr Ile
271          355                   360                   365
274 Gly Pro Glu Met Thr Trp Pro Val Tyr Trp Ser Gly Thr Ile Gly Glu
275      370                   375                   380
278 Ala Lys Lys Asp Val Glu Lys Ala Ala Lys Arg Ile Thr Gln Gln Tyr
279 385                   390                   395                   400
282 Gly Cys Pro Ala Tyr Pro Val Val Ala Lys Ala Leu Val Thr Gln Ala
283              405                   410                   415
286 Ser Ser Ala Ile Pro Val Val Pro Leu Tyr Ile Cys Leu Leu Tyr Arg
287              420                   425                   430
290 Val Met Lys Glu Lys Gly Thr His Glu Gly Cys Ile Glu Gln Met Val
291          435                   440                   445
294 Arg Leu Leu Thr Thr Lys Leu Tyr Pro Glu Asn Gly Ala Pro Ile Val
295      450                   455                   460
```

RAW SEQUENCE LISTING   DATE: 04/26/2006
PATENT APPLICATION: US/10/574,902   TIME: 07:34:04

Input Set : **F:\Final Sequence List-13987-00003-US.txt**
Output Set: **N:\CRF4\04262006\J574902.raw**

```
298 Asp Glu Ala Gly Arg Val Arg Val Asp Asp Trp Glu Met Ala Glu Asp
299 465                     470             475             480
302 Val Gln Gln Ala Val Lys Asp Leu Trp Ser Gln Val Ser Thr Ala Asn
303                 485             490             495
306 Leu Lys Asp Ile Ser Asp Phe Ala Gly Tyr Gln Thr Glu Phe Leu Arg
307         500             505             510
310 Leu Phe Gly Phe Gly Ile Asp Gly Val Asp Tyr Asp Gln Pro Val Asp
311         515             520             525
314 Val Glu Ala Asp Leu Pro Ser Ala Ala Gln Gln
315     530             535
318 <210> SEQ ID NO: 3
319 <211> LENGTH: 837
320 <212> TYPE: DNA
321 <213> ORGANISM: Euglena gracilis
324 <400> SEQUENCE: 3
325 gggtggtata atacggtggc gttcgagaag gccgccctgg aggccgggct gtacgcccgg     60
327 agccttaatg gcgacgcatt cgactctaca acgaaggcgc ggacggttga ggcgatcaag    120
329 cgggatctcg gcacggtgga cctcgtggtg tacagcatcg ccgccccgaa acggacggac    180
331 cctgccaccg gcgtcgtcca caaggcctgc ctgaagccca tcggcgccac ctacaccaac    240
333 cgcactgtga acaccgacaa ggcggaggtg accgatgtca gcatcgagcc ggcctccccc    300
335 gaggagatcg ccgacacggt gaaggtgatg ggcggggagg actgggagct ttggatccag    360
337 gcactgtcgg aggccggcgt gctggcggag ggggccaaga cggtggcata ctcctacatc    420
339 ggccccgaga tgacgtggcc cgtgtactgg tccggcacca ttggggaggc caagaaggac    480
341 gtggagaagg ccgctaagcg catcacacag cagtacggct gcccagcata cccggtggtg    540
343 gccaaggcct tggtcaccca ggccagctct gccatcccgg tggtgccgct ctacatctgc    600
345 ctgctgtacc gcgttatgaa ggagaagggc acccacgagg ctgcatcga gcagatggtg     660
347 cggctgctca ccacgaagct gtaccccggg aacggttccc ccattgtcga tgaggccggg    720
349 cgggtgcggg tggatgactg ggagatggcg gaggatgtgc agcaggctgt gaaggacctc    780
351 tggagccagg tgaacactgc caacctcaag gacatttccg attttgccgg ctatgaa      837
354 <210> SEQ ID NO: 4
355 <211> LENGTH: 1329
356 <212> TYPE: DNA
357 <213> ORGANISM: Euglena gracilis
360 <220> FEATURE:
361 <221> NAME/KEY: CDS
362 <222> LOCATION: (1)..(1329)
365 <400> SEQUENCE: 4
366 atg ggc cgc cgc ggg gcc gca gtg ggc ggc gcc gct cgt ggc cgc ggc      48
367 Met Gly Arg Arg Gly Ala Ala Val Gly Gly Ala Ala Arg Gly Arg Gly
368 1               5                   10                  15
370 ctc ctc cgc act ggc gct gtg gtg gtg aac atc gcc acc gcg gtc acc      96
371 Leu Leu Arg Thr Gly Ala Val Val Val Asn Ile Ala Thr Ala Val Thr
372             20                  25                  30
374 cac ctg gcc ccc ccg atg gcg atg ttc acc acc aca gcg aag gtc atc     144
375 His Leu Ala Pro Pro Met Ala Met Phe Thr Thr Thr Ala Lys Val Ile
376         35                  40                  45
378 cag ccc aag att cgt ggc ttc atc tgc acg acc acc cac ccg atc ggc     192
379 Gln Pro Lys Ile Arg Gly Phe Ile Cys Thr Thr Thr His Pro Ile Gly
380     50                  55                  60
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 04/26/2006
PATENT APPLICATION: **US/10/574,902**   TIME: 07:34:05

Input Set : **F:\Final Sequence List-13987-00003-US.txt**
Output Set: **N:\CRF4\04262006\J574902.raw**

<u>Please Note:</u>
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:12; N Pos. 3,15,18
Seq#:13; N Pos. 9,12
Seq#:26; Xaa Pos. 6,7,8,9
Seq#:29; Xaa Pos. 1,2

VERIFICATION SUMMARY   DATE: 04/26/2006
PATENT APPLICATION: **US/10/574,902**   TIME: 07:34:05

Input Set : **F:\Final Sequence List-13987-00003-US.txt**
Output Set: **N:\CRF4\04262006\J574902.raw**

L:13 M:270 C: Current Application Number differs, Replaced Current Application No
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:1428 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:0
L:1451 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:0
L:1599 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26 after pos.:0
L:1638 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29 after pos.:0